



Further Math with Ali Hashir

Non-Parametric Tests

Topical Questions (CIE 9231)

Compiled from CIE 9231 Further Mathematics past papers (2015–2025), sorted by topic. Each question is numbered for this booklet, with its original paper code and session shown beneath.

About this resource: I teach CIE A Level Further Mathematics, and over the years I found that there simply weren't many well-organised, topic-wise resources available for this subject – it's a niche A Level with relatively few candidates, so it doesn't get much attention from resource makers. These topical booklets are my attempt to fix that: every past-paper question, sorted by topic, with its official mark scheme alongside it.

Help us improve these documents for future students. If you spot a mistake, email us at vectorspacepk@gmail.com.

What kind of mistakes might you find? These booklets are compiled by hand from hundreds of past-paper PDFs. I've checked them as carefully as I can, but I'm only human, and the odd error may have slipped through – so please help me catch what I missed. Things to look out for:

- Spelling or typing mistakes
- Wrong final answers
- A mark scheme that doesn't line up with its question (misaligned pages)
- A mark scheme that belongs to the wrong question entirely
- Pre-2020 questions that are technically out of the current syllabus (the syllabus changed in 2020, but some older questions are still included here for practice)
- Missing, extra, or wrongly-labelled marks (e.g. an M1 or A1 that shouldn't be there, or one that's missing)
- Diagrams or images that are unclear, missing, or don't match the question
- Questions placed under the wrong topic

If you spot any of these (or anything else that looks off), please let me know.

This is **Version 1** of this booklet. As readers point out errors, I'll keep correcting and improving these documents, so I'd recommend checking back on the website from time to time and using the latest version rather than relying on a saved copy.

If you think these documents brought you value, consider following me on social media – it helps more than you'd think.

Q1

9231/41, 9231/42 • May/June 2020 • Question 2

The times, in milliseconds, taken by a computer to perform a certain task were recorded on 10 randomly chosen occasions. The times were as follows.

6.44 6.16 5.62 5.82 6.51 6.62 6.19 6.42 6.34 6.28

It is claimed that the median time to complete the task is 6.4 milliseconds.

- (a) Carry out a Wilcoxon signed-rank test at the 5% significance level to test this claim.
[6]
- (b) State an underlying assumption that is made when using a Wilcoxon signed-rank test.
[1]

Q2

9231/43 • May/June 2020 • Question 6

A biologist is studying the effect of nutrients on the heights to which plants grow. A random sample of 24 similar young plants is divided into two equal groups A and B . The plants in group A are fed with nutrients and water and the plants in group B are given only water. After four weeks, the height, in cm, of each plant is measured and the results are as follows.

Group A	12.3	11.8	12.1	13.2	11.1	10.6	13.8	12.0	12.2	12.4	13.5	13.9
Group B	11.7	10.8	10.9	11.3	11.2	12.6	11.0	10.5	11.9	12.5	10.7	11.6

The biologist decides to carry out a test at the 5% significance level to test whether the nutrients have resulted in an increase in growth.

- (a) She carries out a Wilcoxon rank-sum test. Give a reason why this is an appropriate choice of test. [1]
- (b) Carry out the Wilcoxon rank-sum test for these results. [10]

Q3

9231/41, 9231/43 • October/November 2020 • Question 2

Metal rods produced by a certain factory are claimed to have a median breaking strength of 200 tonnes. For a random sample of 9 rods, the breaking strengths, measured in tonnes, were as follows.

210 186 188 208 184 191 215 198 196

A scientist believes that the median breaking strength of metal rods produced by this factory is less than 200 tonnes.

(a) Use a Wilcoxon signed-rank test, at the 5% significance level, to test whether there is evidence to support the scientist's belief. [6]

(b) Give a reason why a Wilcoxon signed-rank test is preferable to a sign test, when both are valid. [1]

Q4

9231/42 • October/November 2020 • Question 2

A large school is holding an essay competition and each student has submitted an essay. To ensure fairness, each essay is given a mark out of 100 by two different judges. The marks awarded to the essays submitted by a random sample of 12 students are shown in the following table.

Student	A	B	C	D	E	F	G	H	I	J	K	L
Judge 1	62	74	52	48	68	55	56	64	37	70	81	59
Judge 2	65	70	47	49	76	74	67	54	50	77	72	75

(a) One of the students claims that Judge 2 is awarding higher marks than Judge 1.

Carry out a Wilcoxon matched-pairs signed-rank test at the 5% significance level to test whether the data supports the student's claim. [7]

It is discovered later that the marks awarded to student A have been entered incorrectly. In fact, Judge 1 awarded 65 marks and Judge 2 awarded 62 marks.

(b) By considering how this change affects the test statistic, explain why the conclusion of the test carried out in part (a) remains the same. [2]

Q5

9231/41, 9231/42 • May/June 2021 • Question 5

Georgio has designed two new uniforms X and Y for the employees of an airline company. A random sample of 11 employees are each asked to assess each of the two uniforms for practicality and appearance, and to give a total score out of 100. The scores are given in the table.

Employee	A	B	C	D	E	F	G	H	I	J	K
Uniform X	82	74	42	59	60	73	94	98	62	36	50
Uniform Y	78	75	63	56	67	82	99	90	72	48	61

- (a) Give a reason why a Wilcoxon signed-rank test may be more appropriate than a t -test for investigating whether there is any evidence of a preference for one of the uniforms. [1]
- (b) Carry out a Wilcoxon matched-pairs signed-rank test at the 10% significance level. [7]

Q6

9231/43 • May/June 2021 • Question 2

A company is developing a new flavour of chocolate by varying the quantities of the ingredients. A random selection of 9 flavours of chocolate are judged by two tasters who each give marks out of 100 to each flavour of chocolate.

Chocolate	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>
Taster 1	72	86	75	92	98	79	87	60	62
Taster 2	84	72	74	95	85	87	82	75	68

Carry out a Wilcoxon matched-pairs signed-rank test at the 10% significance level to investigate whether, on average, there is a difference between marks awarded by the two tasters. [7]

Q7

9231/41, 9231/43 • October/November 2021 • Question 6

The blood cholesterol levels, measured in suitable units, of a random sample of 11 women and a random sample of 12 men are shown below.

Women	51	55	242	167	152	256	75	137	98	238	235	
Men	311	262	170	302	175	320	220	260	72	351	86	333

Carry out a Wilcoxon rank-sum test, at the 5% significance level, to test whether, on average, there is a difference in cholesterol levels between women and men. [9]

Q8

9231/42 • October/November 2021 • Question 4

Applicants for a particular college take a written test when they attend for interview. There are two different written tests, A and B , and each applicant takes one or the other. The interviewer wants to determine whether the medians of the distribution of marks obtained in the two tests are equal. The marks obtained by a random sample of 8 applicants who took test A and a random sample of 8 applicants who took test B are as follows.

Test A	46	32	29	12	33	18	25	40
Test B	36	28	49	37	48	35	41	31

- (a) Carry out a Wilcoxon rank-sum test at the 5% significance level to determine whether there is a difference in the population median marks obtained in the two tests. [6]

The interviewer considers using the given information to carry out a paired sample t -test to determine whether there is a difference in the population means for the two tests.

- (b) Give two reasons why it is not appropriate to use this test. [2]

Q9

9231/41, 9231/42 • May/June 2022 • Question 6

A teacher at a large college gave a mathematical puzzle to all the students. The median time taken by a random sample of 24 students to complete the puzzle was 18.0 minutes. The students were then given practice in solving puzzles. Two weeks later, the students were given another mathematical puzzle of the same type as the first. The times, in minutes, taken by the random sample of 24 students to complete this puzzle are as follows.

18.2	17.5	16.4	15.1	20.5	26.5	19.2	23.2
17.9	18.8	25.8	19.9	17.7	16.2	17.3	16.6
17.1	20.1	20.3	12.6	16.0	21.4	22.7	18.4

The teacher claims that the practice has not made any difference to the average time taken to complete a puzzle of this type.

Carry out a Wilcoxon signed-rank test, at the 10% significance level, to test whether there is sufficient evidence to reject the teacher's claim. [10]

Q10

9231/43 • May/June 2022 • Question 5

A manager claims that the lengths of the rubber tubes that his company produces have a median of 5.50 cm. The lengths, in cm, of a random sample of 11 tubes produced by this company are as follows.

5.56 5.45 5.47 5.58 5.54 5.52 5.60 5.35 5.59 5.51 5.62

It is required to test at the 10% significance level the null hypothesis that the population median length is 5.50 cm against the alternative hypothesis that the population median length is not equal to 5.50 cm.

Show that both a sign test and a Wilcoxon signed-rank test give the same conclusion and state this conclusion. [9]

Q11

9231/41, 9231/43 • October/November 2022 • Question 3

A large college is holding a piano competition. Each student has played a particular piece of music and two judges have each awarded a mark out of 80. The marks awarded to a random sample of 14 students are shown in the following table.

Student	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Judge 1	79	54	63	74	69	52	50	57	55	42	63	55	56	48
Judge 2	75	62	60	73	76	41	31	51	45	55	49	50	65	36

(a) One of the students claims that on average Judge 1 is awarding higher marks than Judge 2. Carry out a Wilcoxon matched-pairs signed-rank test at the 5% significance level to test whether the data supports the student's claim. [7]

(b) Give a reason why it is preferable to use a Wilcoxon matched-pairs signed-rank test in this situation rather than a paired sample t -test. [1]

Q12

9231/42 • October/November 2022 • Question 6

The manager of a technology company A claims that his employees earn more per year than the employees at technology company B . The amounts earned per year, in hundreds of dollars, by a random sample of 12 employees from company A and an independent random sample of 12 employees from company B are shown below.

Company A	461	482	374	512	415	452	502	427	398	545	612	359
Company B	454	506	491	384	361	443	401	472	414	342	355	437

(a) Carry out a Wilcoxon rank-sum test at the 5% significance level to test whether the manager's claim is supported by the data. [9]

(b) Explain whether a paired sample t -test would be appropriate to test the manager's claim if earnings are normally distributed. [1]

Q13

9231/41, 9231/42 • May/June 2023 • Question 4

A random sample of 13 technology companies is chosen and the numbers of employees in 2018 and in 2022 are recorded.

Company	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>K</i>	<i>L</i>	<i>M</i>
Number in 2018	104	19	126	234	970	514	35	149	429	12	86	304	1104
Number in 2022	106	24	127	228	1012	525	32	156	449	24	78	294	1154

A researcher claims that there has been an increase in the median number of employees at technology companies between 2018 and 2022.

(a) Carry out a Wilcoxon matched-pairs signed-rank test, at the 5% significance level, to test whether the data supports this claim. [7]

The researcher notices that the figures for company *G* have been recorded incorrectly. In fact, the number of employees in 2018 was 32 and the number of employees in 2022 was 35.

(b) Explain, with numerical justification, whether or not the conclusion of the test in part (a) remains the same. [2]

Q14

9231/43 • May/June 2023 • Question 3

A large number of students took two test papers in mathematics. The teacher believes that the marks obtained in Paper 1 will be higher than the marks obtained in Paper 2. She chooses a random sample of 9 students and compares their marks. The marks are shown in the table.

Student	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>
Paper 1	46	73	55	64	86	42	66	68	60
Paper 2	41	66	61	63	90	40	58	42	70

- (a) Carry out a Wilcoxon matched-pairs signed-rank test, at the 5% significance level, to test whether the data supports the teacher's belief. [7]
- (b) State an assumption that you have made in carrying out the test in part (a). [1]

Q15

9231/41, 9231/43 • October/November 2023 • Question 6

A school is conducting an experiment to see whether the distance that children can throw a ball increases in hot weather. On a cold day, all the children at the school were asked to throw a ball as far as possible. The distances thrown were measured and recorded. The median distance thrown by a random sample of 25 of the children was 22.0 m. The children were asked to throw the ball again on a hot day. The distances thrown by the same 25 children were measured and recorded and these distances, in m, are shown below.

21.2	23.5	22.9	18.6	19.4
22.1	26.5	20.2	25.7	20.6
22.3	17.4	22.2	27.0	23.9
28.2	22.6	27.2	23.0	23.7
19.8	22.7	23.3	21.5	24.3

The teacher claims that on average the distances thrown will be further when it is hot.

Carry out a Wilcoxon signed-rank test, at the 5% significance level, to test whether the data supports the teacher's claim. [10]

Q16

9231/42 • October/November 2023 • Question 5 (part a)

A company is deciding which of two machines, X and Y , can make a certain type of electrical component more quickly. The times taken, in minutes, to make one component of this type are recorded for a random sample of 8 components made by machine X and a random sample of 9 components made by machine Y . These times are as follows.

Machine X	4.0	4.6	4.7	4.8	5.0	5.2	5.6	5.8	
Machine Y	4.5	4.9	5.1	5.3	5.4	5.7	5.9	6.3	6.4

The manager claims that on average the time taken by machine X to make one component is less than that taken by machine Y .

- (a) Carry out a Wilcoxon rank-sum test at the 5% significance level to test whether the manager's claim is supported by the data. [6]

Q17

9231/41, 9231/42 • May/June 2024 • Question 2

A large number of students are taking a Physics course. They are assessed by a practical examination and a written examination. The marks out of 100 obtained by a random sample of 15 students in each of the examinations are as follows.

Student	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Practical examination	66	63	24	52	59	76	88	51	48	36	91	72	68	67	60
Written examination	63	57	39	50	47	71	87	65	56	39	78	70	61	62	70

Use a sign test, at the 10% significance level, to test whether, on average, the practical examination marks are higher than the written examination marks. [5]

Q18

9231/41, 9231/42 • May/June 2024 • Question 3

A factory produces metal discs. The manager claims that the diameters of these discs have a median of 22.0mm. The diameters, in mm, of a random sample of 12 discs produced by this factory are as follows.

22.4 20.9 22.8 21.5 23.2 22.9 23.9 21.7 19.8 23.6 22.6 23.0

- (a) Carry out a Wilcoxon signed-rank test, at the 10% significance level, to test whether there is any evidence against the manager's claim. [7]
- (b) State an assumption that is necessary for this test to be valid. [1]

Q19

9231/43 • May/June 2024 • Question 1

A college uses two assessments, X and Y , when interviewing applicants for research posts at the college. These assessments have been used for a large number of applicants this year.

The scores for a random sample of 9 applicants who took assessment X are as follows.

21.4 24.6 25.3 22.7 20.8 21.5 22.9 21.3 22.3

The scores for a random sample of 10 applicants who took assessment Y are as follows.

20.9 23.5 24.8 21.9 23.4 24.0 23.8 24.1 25.1 25.8

The interviewer believes that the population median score from assessment X is lower than the population median score from assessment Y .

Carry out a Wilcoxon rank-sum test, at the 1% significance level, to test whether the interviewer's belief is supported by the data. [7]

Q20

9231/41, 9231/43 • October/November 2024 • Question 2

A school with a large number of students is updating its logo. Each student has designed a new logo and two teachers have each awarded a mark out of 50 for each logo. The marks awarded to a random sample of 12 students are shown in the following table.

Student	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>K</i>	<i>L</i>
Teacher 1	36	38	40	36	22	34	45	44	48	35	28	30
Teacher 2	38	42	32	41	32	41	42	50	36	44	42	41

One of the students claims that Teacher 2 is awarding higher marks than Teacher 1.

- (a) Carry out a Wilcoxon matched-pairs signed-rank test, at the 5% significance level, to test whether the data supports the claim. [7]

It was later discovered that Teacher 1 had entered her mark for student *C* incorrectly. Her intended mark was 24 not 40. This was corrected.

- (b) Determine whether this correction affects the conclusion of the test carried out in part (a). [2]

Q21

9231/42 • October/November 2024 • Question 6

A sports college keeps records of the times taken by students to run one lap of a running track. The population median time taken is 51.0 seconds. After a month of intensive training, a random sample of 22 new students run one lap of the track, giving times, in seconds, as follows.

51.3	52.0	53.4	49.2	49.3	51.1	52.2	47.2
53.0	48.5	49.4	50.3	50.8	51.6	49.1	52.3
51.8	52.4	47.9	48.9	50.6	51.9		

It is claimed that the intensive training has led to a decrease in the median time taken to run one lap of the track.

Carry out a Wilcoxon signed-rank test, at the 5% significance level, to test whether there is sufficient evidence to support the claim. [9]

Q22

9231/41, 9231/42 • May/June 2025 • Question 5

Two jigsaw puzzles have the same number of pieces with identical shapes but have different pictures printed on them. One puzzle has a seaside picture and the other has a cartoon picture. A researcher believes that children will complete the cartoon puzzle more quickly. To test this belief, 10 children are randomly selected. The time taken in seconds for each child to complete each puzzle is recorded below.

Child	A	B	C	D	E	F	G	H	I	J
Seaside	182	130	193	181	192	204	184	192	180	189
Cartoon	161	111	195	159	202	200	168	165	145	160

(a) Carry out a Wilcoxon matched-pairs signed-rank test at the 5% significance level to test the researcher's belief. [6]

(b) Show that using a paired-sample sign test at the 5% significance level would result in the opposite conclusion to that found in part (a). [3]

It was later discovered that the experiment had been conducted such that each child completed the seaside puzzle first followed by the cartoon puzzle.

(c) Comment on the validity of using this experiment to test the researcher's belief. [1]

Q23

9231/43 • May/June 2025 • Question 4

A researcher claims that older people take longer to react to a sudden loud noise than younger people. To investigate this, the researcher randomly selects 6 people over 50 years old and 8 people under 25 years old and records their reaction times, in milliseconds, to a sudden loud noise. The reaction times are as follows.

Over 50	198	212	217	229	235	242			
Under 25	178	181	183	192	203	209	223	231	

Carry out a Wilcoxon rank-sum test at the 5% significance level to test the researcher's claim. [8]

Q24

9231/44 • May/June 2025 • Question 2

The level of sound produced by a particular type of machine was measured for a random sample of 11 such machines. The results, in suitable units, are shown below.

Machine	A	B	C	D	E	F	G	H	I	J	K
Sound level	7.66	8.48	8.21	7.98	8.01	7.77	8.25	8.11	8.03	8.16	7.92

- (a) Use a Wilcoxon signed-rank test to test whether the average sound level produced by this type of machine is more than 8.00. Use a 5% significance level. [6]
- (b) Give a reason why a Wilcoxon signed-rank test may be more appropriate than a t -test in this case. [1]

Q25

9231/41, 9231/43 • October/November 2025 • Question 4

A researcher believes that the median m of a population has changed from its known previous value m_0 . The researcher collects a random sample of size 28. She ranks the data and calculates a test statistic T using the Wilcoxon signed-rank test. The conclusion of the test carried out at a 1% significance level is that there is not sufficient evidence to support her belief.

Using a normal approximation, find the least possible value of T . [5]

Further Math with Ali Hashir
@vectorspacepk • www.vectorspace.pk

Q26

9231/42 • October/November 2025 • Question 1

A large company claims that the median salary of its employees is \$32 500. The salaries (\$) of 15 randomly selected employees are listed below.

18750	30500	125000	42500	25000
26000	52500	23000	27500	19500
25500	33000	30000	21500	29000

(a) Explain why a Wilcoxon signed-rank test may not be appropriate to test the company's claim in this case. [1]

(b) Carry out a sign test at the 10% significance level to investigate the company's claim. [5]

Q27

9231/44 • October/November 2025 • Question 6

Students of the same age from two schools, school A and school B , take a large number of quizzes throughout the year and are each awarded a mark out of 1000. The marks of 123 students in school A and 147 students in school B are ranked from lowest (rank 1) to highest (rank 270). The sum of the ranks of the students from school A is 15355.

- (a) Carry out a Wilcoxon rank-sum test at the 5% significance level to investigate whether there is a difference in average marks between the students in school A and school B . [6]
- (b) State an assumption that is required for the Wilcoxon rank-sum test to be valid. [1]

Help us improve these documents for future students. If you spot a mistake, email us at vectorspacepk@gmail.com.

Written by Ali Hashir.

Answers available in the companion mark scheme booklet.



Instagram: [@vectorspacepk](https://www.instagram.com/vectorspacepk)

Website: www.vectorspace.pk